The Environmental Services Data and Information Management/National Estuarine Research Reserve System Data Rescue Project: Overview and Lessons Learned

A White Paper

by

David Stein
Technology Planning and Management Corporation
National Oceanic and Atmospheric Administration Coastal Services Center
2234 South Hobson Avenue
Charleston, South Carolina 20405

and

Dr. Dwight D. Trueblood
Cooperative Institute for Coastal and
Estuarine Environmental Technology
Room 103 Nesmith Hall
University of New Hampshire
131 Main Street
Durham, New Hampshire 03824-3597

NOAA Technical Memorandum NOS/OCRM/ERD-00-02 and NOS/CSC-/20111-pub

U.S. DEPARTMENT OF COMMERCE
National Oceanic and Atmospheric Administration / National Ocean Service / Office of Ocean and Coastal Resource Management

Table of Contents

1.0 Introduction

In the fall of 1993, 34 United States coastal zone program managers were surveyed by the National Oceanic and Atmospheric Administration (NOAA) to determine their resource management and information needs (NOAA Coastal Committee, 1995). The coastal manager's top priority was to gather information about nonpoint sources of pollution (20 states); preventing wetland habitat loss (17 states); scientific data linking development activity to adverse resource impacts (15 states); managing development impacts to avoid erosion/storm damage (15 states); and mediating multiple use conflicts (13 states). These coastal managers also identified a need for some type of information clearinghouse.

Of the coastal data that exist within the U.S., much is collected by different researchers, in different formats and on different media. In addition, coastal data are often much more complex to manage than other types of data because of the large amount of natural variability that occurs in coastal ecosystems. Good quality coastal data and metadata can lead to an increased overall understanding of coastal ecosystems. However, to be used for analysis and decision making, these data have to be organized, prioritized, and made available for use by researchers and coastal managers.

The Environmental Services Data and Information Management (ESDIM) National Estuarine Research Reserve System (NERRS) Data Rescue project (hereafter called the ESDIM project) was initiated in 1997 to address the need for a broad range of scientifically supported information within our nations' estuaries. The goal of the ESDIM project was to provide access for state coastal zone management programs, and the broader scientific community, to an integrated data-sharing system that contains understandable and useful information.

The project objectives were as follows:

- Rescue, organize, and make easily available all background data for the sites that will help make up part of the national environmental monitoring framework;
- Identify, digitize, and rescue historical U.S. coastal ocean data and metadata that are in immediate danger of being lost, and/or are not readily available to the user community, and make them accessible via the Internet and CD-ROM;
- Provide coastal zone data and information as understandable assessments for policy makers and the public;
- Lead to the development of an integrated, interagency and nongovernmental information base for predicting, assessing, and managing coastal resources for long-term sustainable use.

The data types of greatest interest for this project included: temperature, salinity, oxygen, sediment characteristics, nutrients (nitrate, phosphate, silicate), phytoplankton (primary productivity, biomass, pigments, species), and zooplankton, benthos and fish (secondary productivity, biomass, species), land use changes (via in situ and remote sensing), contaminants, and surface meteorological parameters.

The primary focus of the project was on data rescue. For the purposes of this project, data rescue can be defined as transferring data from an unstable or inaccessible format to a digital format. Data rescue also includes identifying and organizing inaccessible or hard-to-find data, and making those data available to the coastal research and management community. Another important component of the data rescue process was to document all data using the Federal Geographic Data Committee (FGDC) metadata standard.

2.0 ESDIM Data Rescue

The following section provides a chronology of the project, from NERRS site selection to the development of the ESDIM project data clearinghouse. Below is a timeline of the project.

1) Site selection: Winter 1997
2) Hardware and Software Purchases: October 1997 to March 1998
3) Technical Training: January 1998 to July 1998
4) Data Selection and Rescue: April 1998 to March 2000
5) Data Delivery and Management: May 1998 to May 2000

This ESDIM project was a three-year effort administered through existing and ongoing National Ocean Service (NOS) and National Environmental Satellite, Data, and Information Service (NESDIS) programs, contracts, and cooperative agreements. The NOAA Office of Ocean and Coastal Resource Management (OCRM) administered the project while the NOAA Coastal Services Center provided the technical leadership. Project staff consisted of individual contractors and reserve staff that worked under the direction of NOAA principal investigators and on site NERRS research coordinators. Staffing for this project consisted of a project manager from NOAA OCRM who was responsible for budgeting, oversight, and scientific input; a technical coordinator from the NOAA Coastal Services Center who was responsible for geographic information system (GIS) and metadata training, and development and maintenance of the project Internet site; principal investigators at each of the reserves, made up of research coordinators and site mangers, who were responsible for orchestrating and guiding the data rescue at each of the sites; and individual contractors, hired by site principal investigators, who were responsible for implementing the data rescue and restoration at each of the sites. The project principal investigators are listed below.

Dr. Dwight Trueblood, project manager, NOAA Office of Ocean and Coastal Resource Management
Dr. Michael Crosby, project consultant, NOAA/US Agency for International Development
Al Beck, principal investigator, Narragansett NERR
Virginia Lee, principal investigator, University of Rhode Island
Dr. Doug Bulthuis, principal investigator, Padilla Bay NERR
Lee Edmiston, principal investigator, Apalachicola Bay NERR

Carmen Gonzalez, principal investigator, Jobos Bay NERR
Nicholas Schmidt, principal investigator, NOAA Coastal Services Center
David Stein, technical coordinator, Technology Planning and Management Corporation

2.1 Study Sites Selected

The ESDIM project focused on the following four National Estuarine Research Reserve System (NERRS) sites: Apalachicola Bay; Narragansett Bay; Jobos Bay; and Padilla Bay. The sites were selected based on their potential for historic and recent data bases. They were also selected, in part, based on continuation of long-term environmental data gathering, particularly as index sites for the national environmental monitoring framework, and on their geographic distribution.

Apalachicola Bay NERR) Florida

The second largest of the 25 currently existing reserves, the Apalachicola Bay National Estuarine Research Reserve (NERR) includes two barrier islands and a portion of a third, the lower 52 miles of the Apalachicola River and its associated floodplain, portions of adjoining uplands, and the Apalachicola Bay system (Figure 1). The Apalachicola reserve is 55 miles southeast of Panama City, Florida. The overall high water quality of the Apalachicola estuary, with the combined effects of other factors, provide the ideal living conditions for estuarine biota and have resulted in the creation of a highly productive estuarine system. The myriad of habitats found within the reserve support a wide range of plant and animal species, many of which are threatened or endangered.

Figure 1. Apalachicola Bay

Jobos Bay NERR) Puerto Rico

On the southern coast of Puerto Rico (Figure 2), between the municipalities of Salinas and Guayama, approximately 2,883 acres of mangrove forest and freshwater wetlands have been designated as the Jobos Bay National Estuarine Research Reserve. The reserve includes 15 tear-shape islets, known as Cayos Caribe, and the Mar Negro area, which consists of a mangrove forest, and a complex system of lagoons and channels interspersed with salt and mud flats. Freshwater wetlands and subtropical dry forest vegetation are found on the northern island boundaries. Cayos Caribe islets are fringed by coral reefs and seagrass beds, with small beach deposits and upland areas. The Jobos Bay NERR comprises approximately 2,883 acres of land and water divided into two main components: Mar Negro mangrove forest and Cayos Caribe Islets.

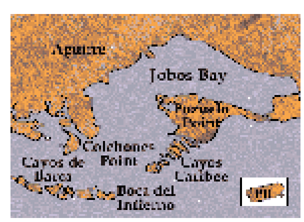

Figure 2. Jobos Bay

Narragansett Bay NERRS) Rhode Island

The Narragansett Bay reserve is located approximately 10 miles southeast of Providence, Rhode Island (Figure 3). The Narragansett Bay National Estuarine Research Reserve is composed of 2,478 acres of land on Prudence, Patience, and Hope Islands; and 1,781 acres of water adjoining the islands. The Narragansett Bay reserve includes undisturbed salt marshes, tidal flats, rock shores, open waters, upland fields, forests, and an historic farm site. The reserve contains the major watershed and largest stream on Prudence Island.

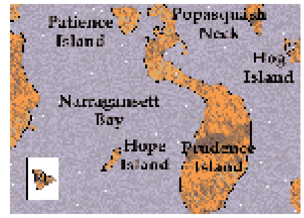

Figure 3. Narragansett Bay

Padilla Bay NERRS) Washington

The Padilla Bay National Estuarine Research Reserve (NERR) is an estuary at the edge of the Skagit River delta (Figure 4). Eight miles long and three miles across, the reserve contains thousands of acres of the ecologically valuable eelgrass, which is habitat to salmon, crab, perch, and herring. Overall, the reserve encompasses 11,000 acres of intertidal and subtidal habitat and 200 acres of upland. Because the bay is filled with sediment from the Skagit River, the bottom is noticeably shallow, flat, and muddy. It is so shallow that almost the whole bay is intertidal. This means that it is flooded at high tide, but when the tide goes out, the whole bay empties out exposing miles and miles of mud flats. This condition allows unusually large eelgrass meadows to grow. There are nearly 8,000 acres of eelgrass in Padilla Bay. Eelgrass is valuable because it is habitat for wildlife and commercially harvested animals. Eelgrass is used as a nursery by salmon, crab, perch, and herring. Eelgrass is also home for millions of worms, shrimp, and clams. Other invertebrates are food for great blue herons, eagles, otters, seals, as well as humans.

Figure 4. Padilla Bay

2.2 Hardware and Software Purchases

To accomplish the goals of the project, desktop GIS systems were purchased for the four sites. The following hardware and software were purchased:

ArcView 3.0a
Dell Optiplex GXPRO
Intel 200Mhz Pentium Pro Processor
2GB Hard Disk Drive
32MB RAM
256K CACHE
TRIOV+ PCI 2MB Video Card
Ultrascan 800HS color monitor(15')
3.5 1.44MB Floppy Disk

12/24X IDE CD ROM
Hewlet Packard ScanJet 4c with OmniPage OCR Software
Zip Drive

In the second year of the project, CD-ROM writers were purchased for the sites so that project data could be distributed to interested parties and final products could be developed. This system configuration met standards developed for the entire NERR system.

2.3 Training

During year one and two of the project, the technical coordinator traveled to each of the sites to conduct site-specific GIS training, and train the contractors on how to implement FGDC metadata for the data they were rescuing. An important component of the project was the development of FGDC-compliant metadata. Software developed by the NOAA Coastal Services Center, known as the ArcView Metadata Collector, was used to collect metadata for GIS files, while standard templates were used for biological and other research data. During year two of the project, three of the four sites sent representatives to attended a National Biological Information Infrastructure (NBII) workshop given by the Biological Resources Division of the US Geological Survey at NOAA's Coastal Services Center. This workshop provided the information and software necessary to document biological and chemical data in compliance with FGDC standards. In addition, principal investigators and contractors from Apalchicola Bay NERR, Jobos Bay NERR, and Padilla Bay NERR attended ArcView training provided by a certified Environmental Systems Research Institute (ESRI) instructor at the NOAA Coastal Services Center. Technical assistance was also given to ESDIM project contractors on an as-needed basis by the project's technical coordinator.

2.4 Data Selection and Rescue

Project staff at each site searched available automated and hard copy coastal databases, catalogs, and inventories to identify data and metadata that were not readily available to the broader management and scientific community and/or were in need of rescue. Some sources of these data included monitoring programs within NOAA, the U.S. Geological Survey (USGS), the Army Corps of Engineers, and the Environmental Protection Agency (EPA) monitoring programs. Other sources included local and regional universities, state and local environmental and natural resource agencies, utility companies, and Non-Governmental Organizations (NGOs). To facilitate prioritization of the data and metadata sets for the rescue and digitization efforts, project staff prepared detailed reports on the results of these searches that included tables by parameter, year, and geographic area.

The methods used to rescue and restore data at each of the sites varied due to staffing resources, technical capacity, and data availability. To guide the overall data rescue process, a project-wide standard operating procedure (SOP) was developed (Appendix A).

Project leaders gave priority to rescuing data sets that were in danger of immediate loss due to deteriorating media. In addition, a major focus was on converting the data not only to digital format, but also to a spatial or geographic information system (GIS) format. These data were also used to develop their individual site GIS. Reports on the results of this search were prepared and sent to the NOAA Coastal Services Center. The reports were also included in the annual project reports. All these files are now accessible through the Internet at the following: *http://www.csc.noaa.gov/pagis/*. Selected data and metadata sets have also been published on CD-ROM. Appendix B includes a complete listing of data rescued and restored during the project.

2.5 Data Delivery, Management, and Dissemination

Early in the project, a central Internet site was developed that provided links to a project description, spatial data, project partners, and on-line mapping. Although site-specific methodologies varied, the method for data documentation (metadata) and delivery to the central Internet site was established early in the project. The FGDC's *Content Standard for Digital Geospatial Metadata* was used to document all data. All data and metadata were then delivered to the project technical coordinator at the NOAA CSC through electronic mail or File Transfer Protocol (FTP). The data and metadata were then quality assured and quality controlled by the Technical Coordinator using the project-wide SOP . If all requirements in the SOP were met, the data and metadata were placed on the project Internet site.

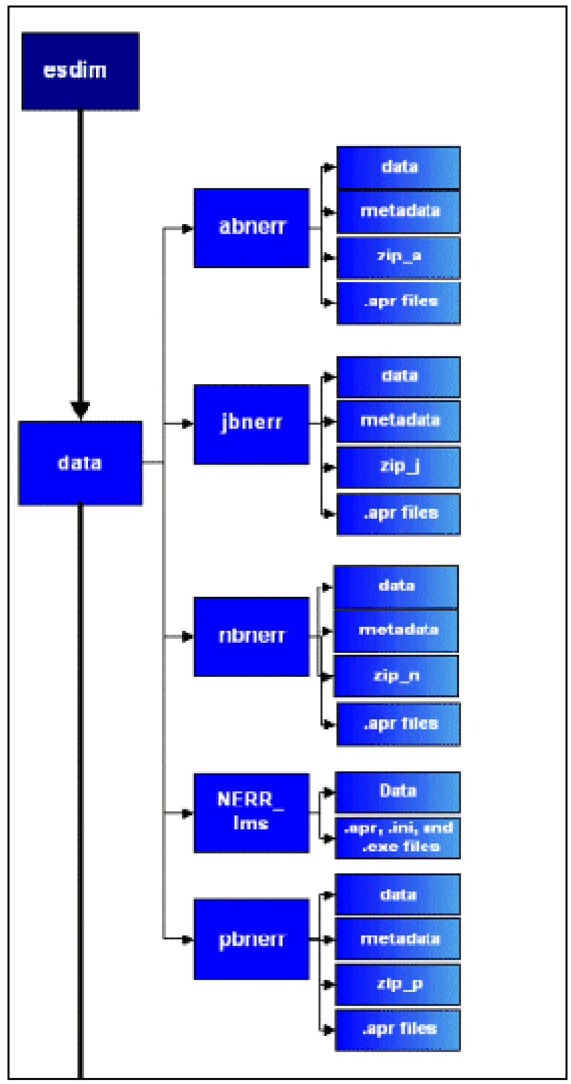

Figure 5. Data Management - Directory Structure

A standardized data management system was also established at the NOAA Coastal Services Center to organize and manage project data. A folder was created for each site. Five sub-folders were then created to store data, metadata, zip files for download, and ArcView files (figure 5). This proved to be a logical and effective data structure.

With regard to final data formats, it was decided that the ArcView Shapefile format would be used for GIS data, and Microsoft Excel would be used for non-GIS data. ArcView Shapefile and Microsoft Excel are widely used data formats, and are also compatible with the hardware and software purchased for the larger Protected Areas Geographic Information System (PAGIS) project. The goal of the ongoing PAGIS project is to build GIS capacity at all of the NERRS and National Marine Sanctuaries (NMS). The GIS data developed during the ESDIM Project was fed directly into the sites' GIS.

With input from the sites, the Internet site was further developed to provide a bulletin board for communication, an on-line mapping application, and a clearinghouse for all rescued data and metadata. Upon completion of the project, the Internet site was merged with the Protected Areas Geographic Information System (PAGIS) project. Since these are compatible projects and PAGIS is an ongoing project, this merger ensured that the ESDIM data would continue to be accessible to the public. The ESDIM Project data clearinghouse can be accessed on the Internet at the PAGIS URL: *http://www.csc.noaa.gov/pagis*. In addition, to reach a wider audience, all project metadata were submitted to the NOAA Coastal Services Center Coastal Information Directory (CID): *http://www.csc.noaa.gov/CID/*. CID is an Internet-based searchable database that contains all of the NOAA Coastal Services Center's project metadata. It is also part of the FGDC's National Geospatial Data Clearinghouse.

3.0 Lessons Learned and Recommendations

An important goal of the ESDIM project was to provide recommendations and identify the types of problems or hurdles the sites encountered while identifying, rescuing, and restoring data. The lessons learned from this project should help other coastal resource managers decide how to allocate resources when conducting similar projects. Since each site approached the project differently, there was some variation in their end products. Collectively, there was a considerable amount of data recovered. The amount of data recovered depended largely on what archives were available to search and on staffing resources. For example, Narragansett Bay NERR had an advantage since the Narragansett National Estuary Program, over the years, had collected a large amount of data around Narragansett Bay. On the contrary, Jobos Bay did not have access to large data archives and, therefore, was unable to rescue as many data sets. Apalachicola Bay utilized a number of data sources in Florida and archives from Florida State University. Padilla Bay focused on identifying and recovering data that could be used in its GIS as well as spatially referencing research studies that had been conducted around Padilla Bay. Padilla Bay staff also spent a considerable amount of time meticulously documenting those data.

Below are the major lessons learned from the perspective of the NOAA project manager and technical coordinator while working on the ESDIM project. Appendices C and D are site-specific lessons learned submitted by Padilla Bay and Apalachicola Bay NERR.

Project-Wide Lessons Learned

1) *Project coordination*
First and foremost, project planning and communication was critical to this project due to geographic and logistical challenges. It is essential that there be good coordination among the project manager, site principal investigators, technical coordinator, and project staff. It is important for project staff to get direct guidance from site coordinators. The goals of the project must be made clear to the project staff at each site so that there are not any questions about what has to be accomplished. It helps to have periodic conference calls to discuss issues and to give updates. It is also vital that the overall project have a person to work with the sites to coordinate and integrate the data rescued and to work with the sites to generate high quality metadata.

2) *Resource Allocation*
Another lesson learned is that it takes time to do this type of work correctly. Sufficient staffing resources need to be allocated to successfully accomplish the goals of a similar project. The first thing that needs to be established is the scope of the project. This will determine staffing and money allocation. For example, this project had part-time site staff and a part-time project-wide technical coordinator. The technical coordinator devoted approximately 25 percent of his time to this project over a three-year period. Most of this time was spent on training, developing, and maintaining the project Internet site, and on administration. While this was an adequate amount of time to spend on the project over a three-year period, it could be more efficient to shorten the length of a similar project and increase the percentage of overall staff time. Another issue that has to be considered is employee turnover. Although not a major problem for the ESDIM project (only one site experienced this), this could have a big impact on the individual site results since time has to be spent training new employees. To hire someone full time with GIS and some scientific experience, the wages have to be competitive. This should be a major consideration when submitting a budget proposal for similar work.

3) *Project Milestones*
At the beginning of the project, quotas for data delivery were set. It was quickly learned that although the project was scheduled for three years, staffing resources, level of experience, and available data determined how many data sets would be recovered. Also, principal investigators working at each of the four NERRs placed different emphasis on the types of data most useful to their reserve. Instead of setting project-wide data delivery milestones, each site developed their own plan. This allowed the sites to focus on quality rather than quantity.

4) *Technical Training*
It is critical that all of the project participants are trained in the technologies (e.g., GIS and metadata creation) that will allow them to meet project goals. This includes principal investigators and site staff. If training is necessary, it should be obtained during the first three months of the project. In addition, if resources such as universities or other regional experts are available, their expertise should be considered. This was an effective strategy for Apalchicola Bay NERR and Narragansett Bay NERR.

Apalachicola Bay NERR utilized resources from Florida State University, while Narragansett Bay NERR utilized resources at the University of Rhode Island.

5) *Metadata*
Metadata are one of the most critical components to any data rescue effort. Well written metadata allow an end user to fully understand data and determine its fitness for use. Since most of the data rescued did not have associated metadata, a considerable amount of time was put into developing metadata. Metadata are equally as imortant as the data; thus, a large percentage of staff time should be devoted to developing fully compliant FGDC metadata.

6) *Internet Connectivity*
Good Internet connectivity is critical to any project that has to communicate through e-mail, FTP or the Web. The Internet provides a fast and efficient way to search for and download data. It is important to have a fast, dedicated line to the Internet so that data can be downloaded efficiently. If a site does not have a dedicated line to the Internet, it is recommended that one be purchased before a project is started or shortly thereafter.

7) *Meetings*
Principal investigator meetings should be held annually. These meetings provide a good venue for addressing critical issues, revisiting project goals, and receiving updates from all parties involved. Two principal investigator meetings were held during the project, both at the Padilla Bay NERR. It is recommended that a meeting agenda be developed and followed to get the most out of these one or two-day meetings. To save travel expenses, the second meeting was held in conjunction with the NERRS research coordinators meeting.

4.0 Conclusion

A wealth of data in analog and outdated formats still exist within our nation's estuaries and coastal environments. As a result, data rescue efforts are becoming increasingly important. The benefits of data rescue include increased data quality, continuity, versatility, and access. Rescued data also have the potential to be used for long-term trend analysis and in geographic information systems. The ESDIM project's primary goal was to rescue and restore data in danger of being lost to deteriorating media, or that were not readily accessible to the coastal resource management community. Although each site focused on data that suited its own resource management needs, the overarching project goals were met. In addition, an equally important component of the project was to learn what it takes to undertake this effort from a technical and programmatic standpoint. The lessons learned from this project should help other sites or individuals when conducting similar work. Project staff at NOAA's Coastal Services Center and the Cooperative Institute for Coastal and Estuarine Environmental Technology (CICEET) are willing to answer specific questions about any component of the project . Questions can be directed to *dave.stein@noaa.gov* or *dwight.trueblood@noaa.gov*.

Appendix A. Project-Wide Standard Operating Procedure

The following standard operating procedure was developed by the project principal investigators to guide the data rescue efforts.

1. Historical and recent data should be evaluated to determine its usability and relevance to research, education, and decision making.

2. If data are deemed worthy of recovery, they should be rescued by scanning, digitizing, reformatting or any other method that will bring them from an unstable or inaccessible format (e.g., hard copy report) to a widely accessible digital format.

3. After the historical data have been converted into digital format, they will be visually cross-checked to assure that the format and quality of the data have not been altered. For example, scanners sometimes mistake the letter *S* for the number *5* and visa versa.

4. Much of the historical data will be in the form of tables or spreadsheets (non-GIS). The final digital format should be in Excel format (.xls).

5. Once the historical data are in a stable format (digital), they should be documented in compliance with Federal Geographic Data Committee (FGDC) metadata standards. Data that are nongeospatial in nature (i.e., meteorological, biological, etc.) should be documented using MetaMaker or a standard template in FGDC format. For geospatial or GIS data that are being rescued, the ArcView Metadata Collector, or any other tool or template that is FGDC compliant, should be used. The ArcView Metadata Collector will work on all computer platforms under ArcView 3.0x.

6. After metadata are developed for data sets, Chew and Spit (CNS) and Metadata Parser (MP) software should be used to verify compliance and to give all metadata the same "look and feel."

7. Historical and recent data that can be spatially referenced should be converted to ArcView Shapefile format in the sites' standardized map projection and/or coordinate system. A critical question is this: "How can we represent something in a GIS that has questionable positional accuracy and/or no coordinate referencing system?" Answer: A point on a paper map, for example, should be transferred to GIS format in a manner that will preserve its positional accuracy as best as possible. The method of transfer will vary from site to site depending on equipment and should be documented accordingly. Below is a generic positional accuracy statement that can go into the metadata record.

> 2.4.1 Horizontal_Positional_Accuracy_Report: This digital data set was derived from original points on a paper map. The horizontal position of the points was preserved as accurately as possible; however, there was no coordinate referencing system on the original maps. See Process_Description (2.5.2.1) in the FGDC *Content Standard for*

Digital Geospatial Metadata for an explanation of the methodology by which the original positional data were transferred into digital format. The purpose of developing this additional data set is so that the geographic position of the research data can be visually displayed and linked together in a GIS.

8. FGDC-compliant metadata should also be generated for the digital positional data that are referencing the original data. This should be a relatively simple process since the "ArcView Metadata Collector" will automatically capture much of the spatial information.

9. Both the original metadata (describing the historical data) and the geographic position metadata (describing the data that spatially reference the historical data) should cross-reference each other. There is a cross reference field (1.14) in the FGDC metadata standard.

10. All restored data and metadata will be made available via a central Internet site, and at each reserve site.

11. The format for distributing GIS files will be ArcView Shapefile (.shp).

Appendix B. Site-Specific Rescued Data) Tables

This appendix includes a complete listing of data rescued and restored during the project.

Apalchicola Bay NERR

Apalachicola Bay searched the data and research archives of many agencies and universities in Florida. The following sources were used for this project:

- Apalchicola Bay National Estuarine Research Reserve (ABNERR)
- Florida Department of Environmental Protection (FDEP)
- National Oceanic and Atmospheric Administration (NOAA)
- U.S. Geologic Survey (USGS)
- Northwest Florida Water Management District
- Florida Fish and Wildlife Conservation Commission
- Florida Marine Research Institute (FMRI)
- Franklin County Planning Office
- U.S. Army Corps of Engineers
- Florida Department of Agriculture and Consumer Services
- Florida State University (FSU) Biology Department
- Various unpublished reports and documents

The table below lists the data set, format, and source of data rescued during this project.

Data Set	Format	Source
Finfish and Shellfish Trawl Stations, Livingston Study, 1973-1983	Shapefile with Graphs	FSU
1995 St. George Island Sea Turtle Nests	Shapefile with database	ABNERR
1996 St. George Island Sea Turtle Nests	Shapefile with database	ABNERR
1997 St. George Island Sea Turtle Nest	Shapefile with database	ABNERR
1998 St. George Island Sea Turtle Nests	Shapefile with database	ABNERR
1996 Dog Island Sea Turtle Nests	Shapefile with database	ABNERR
1997 Dog Island Sea Turtle Nests	Shapefile with database	ABNERR

1998 Dog Island Sea Turtle Nests	Shapefile with database	ABNERR
Coliform Sampling Stations in Apalachicola Bay, 1979 - 1995	Shapefile with database	FDEP
Shorebird Nesting Sites around Apalachicola Bay, 1985 - 1998	Shapefile with database	FDEP
Manatee Sightings in Apalachicola Bay, 1995 - 1998	Shapefile with database	ABNERR
Snapper Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Oyster Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Jacks Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Shrimp Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Blue Crab Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Sea Trout Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Red Drum Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Mullet Landings in Franklin County, 1960 - 1991	Excel Chart with database	FMRI
Listed Fishes, Little Blue Heron, March Wren, Pelican, Priority Wetlands, Red-cockaded Woodpecker, Black Bear, Wading Birds, Snowy Egret, Tricolored Heron, Wakulla Seaside Sparrow, White Ibis, Wood Stork, Wood Stork Foraging Areas	grid file	Florida Game and Fresh Water Fish Commission
Oyster Bars	Shapefile	ABNERR
ANERR Wetland by Class	Shapefile	U.S. Fish and Wildlife Service
ANERR Wetland by Type	Shapefile	U.S. Fish and Wildlife Service
Apalachicola Bay Sediments	Shapefile	Northwest Florida Water Mgmt District (NWFWMD)
Drainage Basins	Shapefile	USGS
Bathymetrry	Shapefile	NOAA

Wetlands (Frankin, Calhoun, Bay, Gadson, Gulf, Jackson, Liberty, Wakulla)	Shapefile	U.S. Fish and Wildlife Service
Calhoun County Wetlands	Shapefile	U.S. Fish and Wildlife Service
Ecoregions	Shapefile	USGS
Seagrass	Shapefile	FMRI
Gulf County Wetlands	Shapefile	U.S. Fish and Wildlife Service
Hydrologic Unit Code Basins	Shapefile	USGS
Wildlife Habitat	Shapefile	Florida Game and Fresh Water Fish Commission
Imagery (Digital Orthophoto Quarter Quads)	.jpg	USGS
Landuse (Frankin, Calhoun, Bay, Gadson, Gulf, Jackson, Liberty, Wakulla)	Shapefile	USGS
ANERR Boundary	Shapefile	ABNERR
Apalachicola River Flow - Blountstown	Shapefile	USGS
Median Household Income	Shapefile	U.S. Census Bureau
Census Tracts	Shapefile	U.S. Census Bureau
Navigation Channels	Shapefile	NOAA
Cities	Shapefile	U.S. Census Bureau/TIGER
Conservation Lands	Shapefile	FDEP
County Areas	Shapefile	U.S. Census Bureau/TIGER
County Lines	Shapefile	U.S. Census Bureau/TIGER
County Seats	Shapefile	U.S. Census Bureau/TIGER
Dredging Information	Shapefile	U.S. Army Corps of Engineers
Ecosystem Management Areas	Shapefile	USGS
Florida Parks	Shapefile	USGS
Lakes	Shapefile	USGS
Marinas	Shapefile	USGS
Navigation Aids	Shapefile	U.S. Coast Guard
Outstanding Florida Waters, Aquatic Preserves	Shapefile	FDEP
Outstanding Florida Waters, Other Categories	Shapefile	FDEP

Outstanding Florida Waters, Special Categories	Shapefile	FDEP
Populated Places	Shapefile	U.S. Census Bureau/TIGER
USGS Topographic Map Index	Shapefile	USGS
Regional Planning Council	Shapefile	FDEP
Streets and Roadways	Shapefile	U.S. Census Bureau/TIGER
Surface Water Classes	Shapefile	FDEP
Water Boundaries	Shapefile	U.S. Census Bureau/TIGER
Coliform and other Water Quality Data	Shapefile	Florida State University

Jobos Bay NERR

The following data sources were used for the Jobos Bay data rescue effort:

- U.S. Geological Survey (USGS)
- University of Puerto Rico
- Jobos Bay National Estuarine Research Reserve (JBNERR)
- NOAA Coastal Services Center

The table below lists the data set, format, and source of data rescued during this project.

Data Set	**Format**	**Source**
Jobos Bay Watershed	Shapefile	USGS
Special Planning Area of JBNERR	Shapefile	JBNERR
Management Area of JBNERR	Shapefile	JBNERR
1995 Landuse/Landcover	Shapefile	University of P.R.
1995 Mangrove Stands	Shapefile	University of P.R.
1977 Mangrove Stands	Shapefile	USGS
1977 Coral Reefs	Shapefile	USGS
The depth of light penetration (seechi disc readings) compared with the depth of the water column	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Summary of variations among three consecutive tows at two stations at Jobos Bay, 1973.	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
A list of the copepod species found in Jobos Bay during 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Fish eggs and larvae in Jobos Bay	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Estimates of total zooplankton biomass in temperate and tropical regions	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total biomass of zooplankton (ml/m3)* in different regions of Jobos Bay, 1973.	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of zooplankton per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of holoplankton per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study

Total number of meroplankton per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of copepods per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of chaetognaths per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of caridean larvae per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of brachyuran zona per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of cirripede nauplii per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Total number of cirripede cypris per m3 in different regions of Jobos Bay, 1973	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Numerical estimates of microzooplankton in Jobos Bay,	Excel Table	Univ. of P.R. - Nuclear Power Plant Study
Biomass and numerical estimates for major metazoan groups other than copepods, no./m3(mm3/m3)	Excel Table	Univ. of P.R. - Nuclear Power Plant Study

Narragansett Bay NERR

The contractors working for Narragansett Bay NERR tapped into data archives around Narragansett Bay. The following data sources were used for the project:

- Narragansett Bay National Estuary Program
- Narragansett Bay National Estuarine Research Reserve (NBNERR)
- Rhode Island Department of Environmental Management (RIDEM)
- University of Rhode Island

The table below lists the data set, format, and source of data rescued during this project.

Data Set	Format	Source
Sub-tidal Habitat Types and Faunal Assemblages	Shapefile	National Estuary Program
Organic Compound Concentrations in Narragansett Bay	Shapefile	National Estuary Program
Sites of Ecological Significance	Shapefile	National Estuary Program
Historical Bathymetric Soundings in Narragansett Bay	Shapefile	National Estuary Program
Shellfish Closure Areas in Narragansett Bay, 1990	Shapefile	National Estuary Program
Shellfish Closure Areas in Narragansett Bay, 1991	Shapefile	National Estuary Program
Narragansett Bay Sediments	Shapefile	National Estuary Program
Sediment Classes in Narragansett Bay, collapsed	Shapefile	National Estuary Program
Hydrolic Partitions of Narragansett Bay	Shapefile	National Estuary Program
Sediment Study Core Sites	Shapefile	National Estuary Program
Nesting Sites and Bird Colonies	Shapefile	National Estuary Program
Major Sources of Nitrogen into Narragansett Bay	Shapefile	National Estuary Program
Major Watersheds Impacting Narragansett Bay	Shapefile	National Estuary Program
Heavy Metal Pollution	Shapefile	National Estuary Program
Narragansett Bay Shellfish Management Areas	Shapefile	National Estuary Program
World Prodigy Degrees of Exposure	Shapefile	National Estuary Program
Hydrologic Sub-Regions of Narragansett Bay	Shapefile	National Estuary Program
Wetland/Salt marsh Monitoring	Shapefile	National Estuary Program
Atmospheric Deposition Study	Shapefile	Univ. of Rhode Island
Boundary of the NBNERR	Shapefile	National Estuary Program

1988 Landuse: Prudence, Patience, Hope Islands	Shapefile	National Estuary Program
Narragansett Bay Basemap	Shapefile	National Estuary Program
Aerial Photograph-Potters Cove, 1997	Tiff image	National Estuary Program
Aerial Photograph-Potters Cove, 1998	Tiff image	National Estuary Program
Aerial Photograph-T-Wharf, 1998	Tiff image	National Estuary Program
CRMC Water Use Catogories	Shapefile	National Estuary Program
RIDEM Water Quality Stations	Shapefile	National Estuary Program
Finfish Sampling Stations in Narragansett Bay	Shapefile	National Estuary Program
HYDROX Sampling Stations	Shapefile	National Estuary Program
Juvenile Finfish Sampling Stations	Shapefile	National Estuary Program
Quahog Sampling Stations in Narragansett Bay	Shapefile	National Estuary Program
Benthic, icthyoplankton, and macrophyte sampling stations	Shapefile	National Estuary Program
Historic Water-Quality Sampling Stations	Shapefile	National Estuary Program
Winter flounder larvae sampling stations for 1990	Shapefile	National Estuary Program
Air Quality Sampling Stations	Shapefile	National Estuary Program
Blackstone Monitoring Stations	Shapefile	National Estuary Program
Cabelli Study Points	Shapefile	National Estuary Program
Historic Chlorophyl Sample Points	Shapefile	National Estuary Program
Historic Coliform Sampling Stations	Shapefile	National Estuary Program
Historic Fish Sample Stations	Shapefile	National Estuary Program
Historic Biological Sampling Stations	Shapefile	National Estuary Program
Historic Phytoplankton Sample Locations	Shapefile	National Estuary Program
Historic Plankton Sample Stations	Shapefile	National Estuary Program
Historic Metal Sampling Stations	Shapefile	National Estuary Program
SQUIRT Stations	Shapefile	National Estuary Program
Wet Weather Sample Locations	Shapefile	National Estuary Program
Fecal Coliform Sampling Stations	Shapefile	National Estuary Program
Smayda Sampling Stations	Shapefile	National Estuary Program
SINBADD (Water Quality) Sampling Stations	Shapefile	National Estuary Program

Padilla Bay NERR

Data rescued at Padilla Bay NERR originated from the following sources:

- Padilla Bay National Estuarine Research Reserve (PBNERR)
- Researchers from the University of Washington
- Puget Sound River Basin Team, Washington State Department of Ecology
- U.S. Geological Survey
- Washington State Department of Ecology
- Skagit County Planning Department
- Western Washington University

The table below lists the data set, format, and source of data rescued during this project.

Data Set	Format	Source
Seagrasses and Macroalgae	Shapefile	PBNERR
Agricultural Use Areas	Shapefile	Washington State Department of Ecology
Buildings with On-site Septic Systems	Shapefile	Washington State Department of Ecology
Drainage Districts	Shapefile	Washington State Department of Ecology
Land Use /Land Cover	Shapefile	Washington State Department of Ecology
Miscellaneous Jurisdiction	Shapefile	Washington State Department of Ecology
Public Land Survey	Shapefile	Washington State Department of Ecology
Proposed Shoreline for Padilla Bay	Shapefile	Washington State Department of Ecology
Proposed Watershed Boundary	Shapefile	Washington State Department of Ecology
Padilla Bay Watershed Zoning	Shapefile	Washington State Department of Ecology
Padilla Bay Wetlands	Shapefile	Washington State Department of Ecology
NOS Bathymetric Survey	Shapefile	NOAA National Ocean Service
Various Water Quality Sampling Sites	Shapefiles	PBNERR
Distribution and Abundance of Fishes Occurring in the Nearshore Surface Waters of Northern Puget Sound, Washington (1974-1976 Sample Sites)	Shapefile	Fresh, K.L , published at the Univ. of Washington
Fish Predation on Dungeness Crab in Padilla Bay, Washington (1987-1988 Sample Sites)	Shapefile	Dinnel, P.A.,et al., Univ. Washington., Fisheries Research Institute

Bacterial Production and Consumption in Microlayer and Subsurface Waters of Padilla Bay, Washington (1994 Sampling Sites)	Shapefile	Thompson, Karen M., Western Washington University
Nutrients and Suspended Solids in Padilla Bay and its Watershed During 1995-96	Shapefile	PBNERR
Nutrient Limitation of Phytoplankton in Padilla Bay, 1992	Shapefile	Bernhard, A.E., Univ. Washington
Opportunistic Feeding habits of the Pacific Staghorn Sculpin (Leptocottus armatus) in Padilla Bay, Washington	Shapefile	Pantalone, V, published at the Univ. of Washington
Padilla Bay Dungeness Crab, Cancer Magister, Habitat Study (1985-1986 Sampling Sites)	Shapefile	Dinnel, P.A.,et al., Univ. Washington., Fisheries Research Institute

Appendix C. Lessons Learned and Recommendations) Submitted by Padilla Bay NERR

Early in the project, the decision was made at Padilla Bay NERR to place an emphasis on development of geospatial data sets. Padilla Bay NERR began this project at the same time that the reserve first acquired desktop GIS capabilities. Thus, staff at Padilla Bay were developing expertise and familiarity with GIS at the same time as the ESDIM project was proceeding. During the course of locating and "rescuing" historical data sets for Padilla Bay, various techniques and strategies were developed. The following are lessons learned about converting nondigital data and converting geospatial data from other organizations to an accessible and useful GIS product.

Padilla Bay NERR has a desktop GIS running ESRI ArcView 3.1 on a Pentium PC and ESRI ArcView 3.0 on a G3 Power Macintosh. Padilla Bay NERR does not have direct hardware links to any larger GIS, such as those of state agencies, universities, or counties. Thus, access to digital data was gained from other organizations through portable disks or via a modem to the Internet.

In general, considerable effort needs to be placed on determining what data will be most useful to each reserve. At Padilla Bay, a high priority was placed on the need to obtain geospatial data from other agencies, and from past studies in ArcView Shapefile format. It was also important for those data to be accessible in a single ArcView project with the same projection and datum.

The time involved in developing complete and accurate metadata for each Shapefile was underestimated. Although this process was time-consuming, an increased appreciation for the value of good metadata was developed since geospatial data from other organizations was obtained with little or no metadata. Decisions about projections, datums, file formats, and directory structure should also be made early in a project as ArcView projects are sensitive to any changes in directory structure or file names.

Data Sources

It was important to consider a broad a range of potential data sources at the start of the project. At Padilla Bay NERR, three sources of data were identified as top priority: hard copy data reports which had sampling sites in or around Padilla Bay; existing geospatial data sets that would provide background for the converted historical datasets; and completion of a digital bathymetry data layer. Potential sources include agency and unpublished historical reports, county and state GIS coverages that can be clipped to the reserve or watershed of the Reserve, raw data in files from historical or recent studies, current monitoring, NOAA digital bathymetry, hard copy historical maps, previous baseline studies, and National Oceanographic Data Center (NODC) data archives. The use of an existing bibliographic database to produce a master list of potential hard copy reports in need of conversion proved useful.

Standard Conventions

Standard conventions for projection, datum, and directory structure proved to be extremely useful.

Without a standard, considerable time would have been spent changing the projection and datum of each Shapefile in order to display these files within one project. Because much of the data were in unprojected geographic coordinates of latitude and longitude, a directory for latitude and longitude and our standard projection was created under each project directory.

<u>Metadata</u>

The NOAA Coastal Services Center's ArcView Metadata Collector was used to compile metadata for each of the datasets. The title, abstract, purpose, and time period of content in each metadata record refer to the Shapefile data set, not the larger work data set. The date of the actual field data collection was stored in the time period information field in the lineage section. It may also be advisable to use a thesaurus for keywords or key places.

The procedures and techniques outlined above include things that were learned during the project. Project staff at Padilla Bay NERR are willing to answer specific questions that other NERRS staff may have in developing similar products. Questions can be directed by e-mail to *sshull@padillabay.gov* or *bulthuis@padillabay.gov*, or by calling (360)-428-1558.

Appendix D. Lessons Learned and Recommendations) Submitted by Apalachicola Bay NERR

Lessons learned from Apalachicola Bay are broken into a novice perspective and an advanced perspective.

Novice Perspective

First, directory structure is important. An ArcView project is essentially a text file containing pathways to the various data files you have included in your project. It is important to thoroughly consider how you wish to organize your data files prior to building an ArcView project. Moving data files after they have been added to an ArcView project will change all the pathways in the project, and much time will be spent redirecting all the files you moved. The ArcScripts section of the ESRI Internet site has a free extension available that will move the files associated with a project to a new location. The extension is called Project File Organizer.

The working directory should also be defined once you have created a new ArcView project. The working directory is where ArcView writes files it creates behind the scenes as you build your project. The default working directory of any new ArcView project is $HOME and is located in the “temp” directory. Make the default working directory a folder that you can easily identify.

Data in Excel spreadsheets can be added to an ArcView project as a table. For point data, latitude and longitude will need to be expressed as decimal degrees and displayed in your table as the fields YCOORD (latitude) and XCOORD (longitude). The Excel file will have to be saved as a .dbf file (dbase IV). Dbase has some limitations, one of which is that column headers will only display the first 10 characters. More information about Dbase is available in the Excel help section. Add the .dbf file per instructions in ArcView for adding tables. Changes made to the .dbf file through Excel will display in the ArcView project each time the project is opened, or when you refresh the table (open the table in Arcview, select Table, then select Refresh). Editing of one field at a time may also be done to the .dbf table while the table is open in ArcView. Conversion of .dbf themes to Shapefiles (.shp) in ArcView allows advanced ArcView editing and display properties to be utilized. However, keep in mind that once a .dbf file is converted to a Shapefile it will not accept changes automatically from Excel. Make sure the data in the .dbf file is in its final form before you convert to a Shapefile. Data in a Shapefile format can be changed or edited but it has to be done one field at a time and is a more cumbersome process than editing the .dbf file through Excel.

Another helpful suggestion is that the “select feature” button on the view toolbar is effective for selecting features that may require editing. After selecting features on the view, open the table for the feature’s theme and select the “promote” button and your record for that selected feature will appear at the top of the table. Edits to the feature’s record can be made this way. Select “refresh” from the Table menu, to make the edit take effect on your view. A Query can also be performed in an open table to edit numerous features. Refer to the ArcView help section for more information on the use of the Query feature.

Project size is another issue that has to be considered. An advantage to having a big and inclusive ArcView project is that many views can be created and organized in a logical manner. It is also easy to copy and paste from one view to another. However, if a big project crashes, a lot of data will be lost and much time will have to spent rebuilding the project.

The following are some things to consider when using the ArcView Metadata Collector. The Metadata Collector Tool collects metadata for the individual themes in an ArcView project. As a result, each individual data layer represented as a theme in ArcView will need to have its own unique metadata written. This should be considered when deciding how to break down or categorize related data. Text written in the Collector program is saved and specific portions can be retrieved for use in metadata on similar data, thus saving some data entry time. The Metadata Collector Tool developed by NOAA's Coastal Services Center is FGDC compliant, works well, and streamlines the process of metadata creation.

There are many on-line resources with free downloadable extensions and scripts that can increase the versatility and scope of a project. A good source to use is the Environmental Systems Research Institutes (ESRI) URL: *http://www.esri.com*.

Advanced Perspective

The real power in an ArcView GIS project lies not so much in developing a large number of fixed views, but rather in the assembly of resources including Shapefiles, data files, and supporting documentation. ArcView projects will not be static entities. Data collection within the reserves will be an ongoing effort for many thematic coverages and the integration of new data will require an expenditure of time and money. The same is true for the various supporting coverages obtained from external agencies, who will release periodic updates to most of the data themes which have been included in the current project. Reserves will need to plan for an investment in project maintenance and updates. The software platform for the study (ArcView) is itself undergoing evolution and Reserves will need to be prepared for both the hardware and software upgrades that will be required to support future versions of the system.

Another issue to consider is that GIS source data come in a wide variety of scales, projections, datums, and resolutions. The source data do not always fit together, and sometimes large amounts of work are needed to match different coverages. Fortunately, many agencies and data clearinghouses perform a lot of this work. The GIS field is so new and active that there is sometimes high turnover in personnel at agencies responsible for certain data. This can pose problems when attempting to trace lineage for some data themes.

Finally, GIS projects can relate tremendous quantities of disparate kinds of data, in theme (data subject), form (both maps and pictures of various kinds), and extent (geographic space). GIS projects offer a perspective on information that affords planning and management personnel levels of insight and understanding that they have never had before.

Appendix E. List of Acronyms

COE - Army Corps of Engineers
CICEET - Cooperative Institute for Coastal and Estuarine Environmental Technology
CSC - Coastal Services Center
ESRI- Environmental Systems Research Institute
EPA - Environmental Protection Agency
ESDIM - Environmental Systems Data and Information Management
FDEP - Florida Department of Environmental Protection
FGDC - Federal Geographic Data Committee
NBII - National Biological Information Infrastructure
NERR - National Estuarine Research Reserve
NESDIS- National Environmental Satellite, Data, and Information Service
NGO - Non-Governmental Organization
NOAA - National Oceanic and Atmospheric Administration
NOS - National Ocean Service
OCRM - Ocean and Coastal Resource Management
SOP - Standard Operating Procedure
TIGER - Topologically Integrated Geographic Encoding and Referencing system
USGS- United States Geologic Survey